SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: SAGAWA, HIROAKI
                UENO, HARUMI
                OSHIMA, ATSUSHI
                KATO, IKUNOSHIN

    (ii) TITLE OF INVENTION: PLASMID

   (iii) NUMBER OF SEQUENCES: 33

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
        (B) STREET: PO BOX 747
        (C) CITY: FALLS CHURCH
        (D) STATE: VA
        (E) COUNTRY: USA
        (F) ZIP: 22040-0747

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

  (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: WEINER, MARC S.
        (B) REGISTRATION NUMBER: 32,181
        (C) REFERENCE/DOCKET NUMBER: 1422-0319P

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 703-205-8000
        (B) TELEFAX: 703-205-8050

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 295 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "2=Val or Leu"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Xaa Pro Leu Asp Lys Asp Leu Gln Lys Ala Lys Ile Ser Ile Thr
1               5                   10              15

Asp Phe Phe Glu Ile Thr Asn Arg Val Leu Asp Tyr Phe Pro Asn Val
            20                  25              30

Ile Asn Asn Thr Val Glu Lys Gly Asp Tyr Leu Ile Ser Ser Ser Asn
        35                  40                  45

Ile Ala Gly Thr Ile Lys Phe Leu Arg Pro Ile Asn Arg Lys Leu Phe
    50                  55                  60

Ile Gln Glu Lys Lys Val Phe Asn Asp Tyr Phe Gln Lys Leu Ile Ile
65              70                  75                      80

Val Phe Glu Asn Ile Arg Asn Lys Lys Thr Val Thr Glu Glu Asp Lys
                85                  90                  95

Ile Ile Ile Asp Arg Val Ile Tyr Thr Ile Gln Gln Ser Ile Gly Ile
            100                 105             110

Gly Leu Asp Leu Met Val Asn Gln Asn Ser Ala Arg Lys His Val Gly
        115                 120                 125

Asn Arg Phe Glu Glu Leu Ile Arg Val Ile Phe Thr Glu Ile Ser Val
    130                 135                 140

Ser Asn Lys Arg Thr Val Leu Gln Ile Pro Tyr Glu Thr Asp Glu Gly
145                 150                 155                 160

Gln Lys Ile Tyr Lys Cys Glu Asn Asp Leu Ile Ile Ser Pro Phe Glu
            165                 170                 175

Asn Val Glu Ser Thr Asn Lys His Leu Asp Glu Asn Glu Ile Val Val
        180                 185                 190

Ser Ile Lys Thr Thr Ser Lys Asp Arg Met Gly Lys Met Phe Ile Asp
    195                 200                 205

Lys Ile Leu Leu Glu Arg Phe Val Lys His Pro Gln Lys Val Ile Gly
    210                 215                 220

Ile Phe Leu Asn Asp Val Gln Arg Lys Glu Asp Asn Asn Ile Ser Phe
225                 230                 235                 240

Thr Leu Val Ser Gly Leu Phe Met Val Tyr Thr Lys Phe Leu Thr Thr
            245                 250                 255

Leu Glu Gly Ile Tyr Tyr Leu Asp Pro Pro Pro Asn Ala Leu Lys Leu
            260                 265                 270

Pro Tyr Ser Asn His Met Lys Arg Phe Ser Asp Leu Ile Thr Glu Asp
        275                 280                 285

Leu Glu Lys Leu Phe Ser Ser
    290                 295
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:

```
        (A)  LENGTH: ●●● base pairs
        (B)  TYPE: nucleic acid
        (C)  STRANDEDNESS: double
        (D)  TOPOLOGY: linear

  (ii)  MOLECULE TYPE: DNA (genomic)



  (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGSTACCAC TGGATAAAGA TTTACAAAAA GCAAAGATTT CAATTACTGA TTTTTTTGAA      60

ATTACAAATA GAGTTTTAGA TTATTTCCCC AATGTAATCA ATAATACAGT TGAAAAAGGA     120

GATTATTTAA TATCCTCATC AAATATTGCT GGAACAATAA AATTCCTAAG ACCAATCAAT     180

AGAAAGTTAT TTATTCAGGA AAAAAAAGTT TTCAATGATT ATTTTCAAAA ACTGATTATA     240

GTTTTTGAAA ATATAAGGAA CAAAAAAACT GTAACAGAGG AAGATAAAAT TATTATTGAT     300

AGGGTAATTT ACACAATACA GCAATCTATT GGAATTGGTT TAGATTTAAT GGTTAATCAA     360

AATAGTGCTA GAAAGCACGT TGGTAACCGA TTTGAAGAAT TAATTAGAGT CATTTTTACA     420

GAAATATCAG TATCGAATAA AAGAACTGTA TTACAAATTC CATATGAAAC TGATGAAGGA     480

CAGAAAATTT ACAAATGCGA GAATGACCTC ATTATTTCTC CTTTTGAAAA TGTAGAATCT     540

ACAAACAAAC ATCTAGATGA AAATGAGATT GTTGTTTCAA TAAAGACAAC ATCAAAAGAT     600

AGGATGGGAA AAATGTTTAT AGATAAAATT TTACTTGAAA GGTTTGTTAA ACACCCTCAA     660

AAAGTTATAG GGATTTTCCT CAATGATGTA CAAAGAAAAG AAGACAACAA TATCAGCTTT     720

ACACTTGTTT CAGGATTATT TATGGTGTAT ACTAAATTCT TAACTACTCT TGAAGGGATC     780

TATTATTTAG ATCCACCACC TAATGCATTG AAACTACCAT ATTCTAATCA TATGAAAAGA     840

TTTTCAGATT TAATTACAGA AGACCTTGAA AAATTATTCT CCTCT                     885
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 215 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TATGGATATG TTCATAAACA CGCATGTAGG CAGATAGATC TTTGGTTGTG AATCGCAACC      60

AGTGGCCTTA TGGCAGGAGC CGCGGATCAC CTACCATCCC TAATGACCTG CAGGCATGCA     120
```

AGCTTGCATG CCTGCAGGTC ATTAGGTACG GCAGGTGTGC TCGAGGCGAA GGAGTGCCTG   180

CATGCGTTTC TCCTTGGCTT TTTTCCTCTG GGACA   215

(2) INFORMATION FOR SEQ ID NO:4:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 215 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TATGTCCCAG AGGAAAAAAG CCAAGGAGAA ACGCATGCAG GCACTCCTTC GCCTCGAGCA   60

CACCTGCCGT ACCTAATGAC CTGCAGGCAT GCAAGCTTGC ATGCCTGCAG GTCATTAGGG   120

ATGGTAGGTG ATCCGCGGCT CCTGCCATAA GGCCACTGGT TGCGATTCAC AACCAAAGAT   180

CTATCTGCCT ACATGCGTGT TTATGAACAT ATCCA   215

(2) INFORMATION FOR SEQ ID NO:5:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 28 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGATCTAGAG CAAACAAAAA AACCACCG   28

(2) INFORMATION FOR SEQ ID NO:6:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGTCTAGATC CCAGAGGAAA AAAG   24

(2) INFORMATION FOR SEQ ID NO:7:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 100 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCGAGATTT AGGTGACACT ATAGAATACG GAATTGTGAG CGGATAACAA TTCCAAGCTT   60

CACAGGAAAC AGACCATGGC TTAAGTAACT AGTGAATTCG   100

(2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 100 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGAATTCACT AGTTACTTAA GCCATGGTCT GTTTCCTGTG AAGCTTGGAA TTGTTATCCG   60

CTCACAATTC CGTATTCTAT AGTGTCACCT AAATCTCGAG   100

(2) INFORMATION FOR SEQ ID NO:9:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AATCCCATGG AACGCTACGA ATCTCTG   27

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGGCCATGG TTATTTTTGA CACCAGACC   29

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TAACTTGAAT CCATGGGTTC TCACCG   26

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TACTCAGTAG CCATGGCTCT CATAGACCG   29

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 308 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Asn Glu Ile Ala Phe Asp Asn Tyr Ser Tyr Ile Pro Lys Leu Lys
1               5                   10              15

Leu Tyr Ser Glu Ile Glu Leu Lys Pro Phe Phe Ile Ser Lys Asn Gly
            20                  25              30

Ser Leu Phe Asn Val Asp Ala Ile Asp Phe Leu Arg Lys Leu Glu Ser
        35                  40                  45

Asn Ser Val Asp Leu Ile Phe Ala Asp Pro Pro Tyr Asn Ile Lys Lys
    50                  55                  60

Ala Glu Trp Asp Ile Phe Ser Ser Gln Asn Glu Tyr Leu Glu Trp Ser
65                  70                  75                  80

Lys Glu Trp Ile Met Glu Ala His Arg Val Leu Lys Asp Asn Gly Ser
                85              90                  95

Leu Tyr Val Cys Gly Phe Ser Glu Ile Leu Ala Asp Ile Lys Phe Ile
            100                 105                 110

Thr Ser Lys Tyr Phe His Ser Cys Lys Trp Leu Ile Trp Phe Tyr Arg
        115                 120                 125

Asn Lys Ala Asn Leu Gly Lys Asp Trp Gly Arg Ser His Glu Ser Ile
    130                 135                 140

Leu Leu Leu Arg Lys Ser Lys Asn Phe Ile Phe Asn Ile Asp Glu Ala
145                 150                 155                 160

Arg Ile Pro Tyr Asn Glu His Thr Val Lys Tyr Pro Gln Arg Thr Gln
                165                 170                 175

Ala Glu Ser Ser Gln Tyr Ser Asn Ser Lys Lys Gln Tyr Ile Trp Glu
            180                 185                 190

Pro Asn Pro Leu Gly Ala Lys Pro Lys Asp Val Leu Glu Ile Pro Thr
        195                 200                 205

Ile Ser Asn Gly Ser Trp Glu Arg Ser Ile His Pro Thr Gln Lys Pro
    210                 215                 220

Val Glu Leu Leu Lys Lys Ile Ile Leu Ser Ser Ser Asn Lys Asp Ser
225                 230                 235                 240

Leu Ile Leu Asp Pro Phe Gly Gly Ser Gly Thr Thr Tyr Ala Val Ala
                245                 250                 255

Glu Ala Phe Gly Arg Lys Trp Ile Gly Thr Glu Leu Asp Lys Asn Tyr
            260                 265                 270

Cys Leu Glu Ile Gln Lys Arg Leu Lys Asp Glu Ser Met Ile Asn Arg
```

88

```
Ile Phe Ser Gly Asp Asp Asp Ser Asn Ser Gln Asn Arg Arg Lys Lys
    290                 295                 300

Leu Arg Gly Glu
305
```

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 924 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
GTGAATGAAA TAGCGTTTGA TAATTACAGT TATATACCAA AATTAAAACT TTATTCGGAA      60

ATCGAGCTTA AACCATTTTT TATTTCAAAA AACGGTTCAC TTTTCAATGT TGATGCTATT     120

GATTTTTTAA GAAAATTAGA GAGTAATTCT GTGGATTTAA TTTTTGCAGA TCCACCTTAT     180

AACATTAAAA AGGCAGAGTG GGATATTTTT TCTTCTCAAA ATGAATATCT CGAATGGAGT     240

AAAGAATGGA TAATGGAAGC TCATAGAGTT TTAAAAGATA ATGGCAGTTT ATATGTTTGT     300

GGCTTTTCAG AAATTCTGGC AGACATAAAA TTTATCACTT CAAAATATTT TCACAGTTGT     360

AAATGGTTGA TTTGGTTCTA TAGAAACAAG GCAAATTTAG GTAAAGATTG GGGACGTTCA     420

CACGAAAGTA TACTGTTATT AAGAAAATCT AAAAATTTTA TTTTTAATAT TGATGAGGCA     480

CGAATCCCGT ATAATGAGCA TACAGTTAAA TATCCACAAA GAACCCAGGC CGAATCTTCG     540

CAATATTCGA ACTCAAAAAA GCAATATATT TGGGAGCCAA ACCCATTAGG AGCTAAGCCA     600

AAAGATGTTT TGGAGATTCC CACAATTTCA AATGGTTCTT GGGAAGAAG TATTCACCCT     660

ACGCAAAAGC CAGTAGAATT GCTTAAAAAA ATAATTTAT CTTCATCTAA TAAAGATAGT     720

TTAATTCTTG ATCCATTTGG TGGTTCGGGA ACTACATATG CTGTTGCGGA AGCTTTTGGC     780

AGAAAATGGA TTGGAACAGA GTTAGATAAA AATTATTGTC TGGAAATTCA AAAGCGATTG     840

AAAGACGAAA GTATGATCAA CAGGATTTTT TCAGGCGATG ATGATTCAAA TTCTCAAAAT     900

AGAAGAAAAA AATTAAGAGG AGAA                                          924
```

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TCGAGATTTA GGTGACACTA TAGAATACA   29

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AGCTTGTATT CTATAGTGTC ACCTAAATC   29

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 54 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCGAGATTTA GGTGACACTA TAGAATACGG AATTGTGAGC GGATAACAAT TCCA   54

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 54 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AGCTTGGAAT TGTTATCCGC TCACAATTCC GTATTCTATA GTGTCACCTA AATC   54

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Leu Pro Leu Asp Lys Asp Leu Gln Lys Ala Lys Ile Ser Ile Thr
1   5   10   15

Asp Phe Phe Glu
            20

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:
        (A) NAME/KEY: primer_bind
        (B) LOCATION: 1..23
        (D) OTHER INFORMATION: /note= "6, 9, 12 = inosine"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATGTTNCCNY TNGAYAARGA YYT   23

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (ix) FEATURE:

(A) NAME/KE█ primer_bind
        (B) LOCATION: 1..23
        (D) OTHER INFORMATION: /note= "9 = inosine"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAGGATTTNC ARAARGCNAA RAT   23

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TAAATCTAAA CCAATTCCAA TAGATTGCTG   30

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TAAATCTAAA CCAATTCCAA TAGATTGCTG   30

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "synthetic DNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAACTGTATT ACAAATTCCA ████TGAAACTG   30

(2) INFORMATION FOR SEQ ID NO:25:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GACAGAAAAT TTACAAATGC GAGAATGACC   30

(2) INFORMATION FOR SEQ ID NO:26:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 1588 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CCATGGCACA CGTTTCAAAA AAGAAATCCT CGAAGTCAAA TATGATGAGA AAAACATCTC   60

AGACATCCTG CATATGACGG TGGATGAAGC ATTGGAATTT TTCTCGGAAA ATCACGAAGA   120

AAAAATTGTA ACCAAACTAA AACCTTTGCA GGACGTTGGT TTGGGTTATC TTCAGTTAGG   180

CCAGTCCTCC TCTACTCTTT CCGGCGGTGA AGCCCAAAGA GTGAAGCTCG CCTCTTTCCT   240

TGTGAAAGGT GTAACGACGG AAAAAACGTT ATTTGTTTTT GATGAACCAT CAACAGGATT   300

ACATTCCAC GACATTCAAA AATTACTGAA ATCACTTCAG GCACTGATAG AATTAGGGCA   360

TTCGGTTGTA GTGATTGAGC ATCAGCCGGA TATTATCAAA TGCGCCGATT ACATCATCGA   420

TGTCGGACCC AATGCCGGAA AATACGGTGG CGAAATTGTT TTCACAGGAA CTCCGGAAGA   480

TTTGGTAAAA GAGAAAAAGT CGTTTACAGG GAAGTATATT AAGGAGAAGT TAAAGTAATT   540

TATTTATATT TGAAGTTATG CTACCACTGG ATAAAGATTT ACAAAAAGCA AAGATTTCAA   600

TTACTGATTT TTTTGAAATT ACAAATAGAG TTTTAGATTA TTTCCCCAAT GTAATCAATA   660

ATACAGTTGA AAAAGGAGAT TATTTAATAT CCTCATCAAA TATTGCTGGA ACAATAAAAT   720

TCCTAAGACC AATCAATAGA AAGTTATTTA TTCAGGAAAA AAAAGTTTTC AATGATTATT   780

93

TTCAAAAACT GATTATAGTT TTTGAAAATA TAAGGAACAA AAAAACTGTA ACAGAGGAAG   840

ATAAAATTAT TATTGATAGG GTAATTTACA CAATACAGCA ATCTATTGGA ATTGGTTTAG   900

ATTTAATGGT TAATCAAAAT AGTGCTAGAA AGCACGTTGG TAACCGATTT GAAGAATTAA   960

TTAGAGTCAT TTTTACAGAA ATATCAGTAT CGAATAAAAG AACTGTATTA CAAATTCCAT   1020

ATGAAACTGA TGAAGGACAG AAAATTTACA AATGCGAGAA TGACCTCATT ATTTCTCCTT   1080

TTGAAAATGT AGAATCTACA AACAAACATC TAGATGAAAA TGAGATTGTT GTTTCAATAA   1140

AGACAACATC AAAAGATAGG ATGGGAAAAA TGTTTATAGA TAAAATTTTA CTTGAAAGGT   1200

TTGTTAAACA CCCTCAAAAA GTTATAGGGA TTTTCCTCAA TGATGTACAA AGAAAAGAAG   1260

ACAACAATAT CAGCTTTACA CTTGTTTCAG GATTATTTAT GGTGTATACT AAATTCTTAA   1320

CTACTCTTGA AGGGATCTAT TATTTAGATC CACCACCTAA TGCATTGAAA CTACCATATT   1380

CTAATCATAT GAAAAGATTT TCAGATTTAA TTACAGAAGA CCTTGAAAAA TTATTCTCCT   1440

CTTAATTTTT TTCTTCTATT TTGAGAATTT GAATCATCAT CGCCTGAAAA AATCCTGTTG   1500

ATCATACTTT CGTCTTTCAA TCGCTTTTGA ATTTCCAGAC AATAATTTTT ATCTAACTCT   1560

GTTCCAATCC ATTTTCTGCC AAAAGCTT   1588

(2) INFORMATION FOR SEQ ID NO:27:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 25 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATATTTGAAG CCATGGTACC ACTGG   25

(2) INFORMATION FOR SEQ ID NO:28:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 26 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "synthetic DNA"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

(2) INFORMATION FOR SEQ ID NO:29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4146 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
AGATCTGGTC ATCCCAAACA AAAATCTTTC GGTTTACGAA GATGCAGTCG CTTCCTGGAA        60

AGGCGAAAGT ATGAGCGAAT GGAAAAAGGA ATTCATCAAA AAAGCCAAAC ATTTCCCAAT       120

TCACAAGCCT TATCATCAAC TCACAAAAGA GCAGAAACAG TTCCTTTGGA AAGGCGATAA       180

AACCAGAAGT TTCCCAAGTA TTGATAATTT TTTCAAAATG CTTGAAGAGA ATCTTTACAA       240

GATCCAATAC CGCGTAATGC TTTCGCGCTA TCGTGGGAAA ACACTTTGCC CCGATTGCGA       300

AGGATTACGA TTGCGGGAAG AAACAAGCTG GGTGAAGATT GACGGACACA ACATTCAGTC       360

TTTGATTGAA TTACCTTTGG ATGAACTCCT GCCATTGATC AAAAGCTTAA AACTGAACGT       420

CCACGACAGA GAAATTGCCA AACGCCTGAC TTACGAAATC GAAACGAGAT TAGAATTCCT       480

GACGAAAGTC GGCCTTGGAT ATCTGACTTT GAACCGAACA TCCAACACGC TTTCCGGAGG       540

AGAAAGCCAG AGAATCAATC TGGCGACAGC TTGGGAAGTT CGCTGGTTGG TTCTATTTAT       600

ATTTTGGATG AGCCGAGCAT TGGTCTGCAT TCCCGCGATA CAGAAAATCT GATTGGTGTC       660

CTCAAACAAC TCCGCGATTT GGGAANTACC GTGATTGTTG TAGAACACGA CGAAGATGTG       720

ATGCTTGCGG CAGNTTACAT TATAGATATT GGCCCNGNAG CGGGCTACCT TGGTGGCGAT       780

CTTGTTTTCA GCGNGGATTA TAAAGAGATG CTGAAGTNTN ATACTTTAAC CGCAAAATAC       840

CTGAATGGCG AACTGAAAAT AGAAGTTCCT GAAAAACGAA GAAAACCGAA GGAATTCATC       900

GCAATAAAAG GTGCCCGCCA GAATAATTTA AAAAATATTG ACGTTGATGT TCCGTTAGAA       960

TGTCTGACAG TTATCACAGG CGTTTCTGGA AGCGGGAAAT CCACTTTGAT GAAGGAAGTG      1020

ATGACCAATG CCATCCAGAT CCAACTGGGA ATGGGCGGCA AAAAGCCGA TTACGATTCG      1080

GTGGAATTCC CGAAAAAGCT GATCCAGAAT ATCGAACTGA TTGACCAGAA CCCAATCGGG      1140

AAATCGTCCC GCTCCAACCC CGTGACTTAT CTGAAAGCTT ACGACGATAT CCGGGATCTT      1200

TTTGCGAAAC AAAAATCCGC AAAAATCCAG GGTTACAAAC CGAAGCATTT CTCCTTCAAT      1260

GTGGATGGCG GAAGATGTGA CGAGTGCAAA GGCGAAGGTA TCATTACCGT ATCAATGCAG      1320
```

```
TTTATGGCGG ACATCGAGCT GGAGTGTGAG CATTGCCATG GCACACGTTT CAAAAAAGAA    1380

ATCCTCGAAG TCAAATATGA TGAGAAAAAC ATCTCAGACA TCCTGCATAT GACGGTGGAT    1440

GAAGCATTGG AATTTTTCTC GGAAAATCAC GAAGAAAAAA TTGTAACCAA ACTAAAACCT    1500

TTGCAGGACG TTGGTTTGGG TTATCTTCAG TTAGGCCAGT CCTCCTCTAC TCTTTCCGGC    1560

GGTGAAGCCC AAAGAGTGAA GCTCGCCTCT TTCCTTGTGA AAGGTGTAAC GACGGAAAAA    1620

ACGTTATTTG TTTTTGATGA ACCATCAACA GGATTACATT TCCACGACAT TCAAAAATTA    1680

CTGAAATCAC TTCAGGCACT GATAGAATTA GGGCATTCGG TTGTAGTGAT TGAGCATCAG    1740

CCGGATATTA TCAAATGCGC CGATTACATC ATCGATGTCG ACCCAATGC CGGAAAATAC    1800

GGTGGCGAAA TTGTTTTCAC AGGAACTCCG GAAGATTTGG TAAAAGAGAA AAAGTCGTTT    1860

ACAGGGAAGT ATATTAAGGA GAAGTTAAAG TAATTTATTT ATATTTGAAG TTATGCTACC    1920

ACTGGATAAA GATTTACAAA AAGCAAAGAT TTCAATTACT GATTTTTTTG AAATTACAAA    1980

TAGAGTTTTA GATTATTTCC CCAATGTAAT CAATAATACA GTTGAAAAAG GAGATTATTT    2040

AATATCCTCA TCAAATATTG CTGGAACAAT AAAATTCCTA AGACCAATCA ATAGAAAGTT    2100

ATTTATTCAG GAAAAAAAAG TTTTCAATGA TTATTTTCAA AAACTGATTA TAGTTTTTGA    2160

AAATATAAGG AACAAAAAAA CTGTAACAGA GGAAGATAAA ATTATTATTG ATAGGGTAAT    2220

TTACACAATA CAGCAATCTA TTGGAATTGG TTTAGATTTA ATGGTTAATC AAAATAGTGC    2280

TAGAAAGCAC GTTGGTAACC GATTTGAAGA ATTAATTAGA GTCATTTTTA CAGAAATATC    2340

AGTATCGAAT AAAAGAACTG TATTACAAAT TCCATATGAA ACTGATGAAG GACAGAAAAT    2400

TTACAAATGC GAGAATGACC TCATTATTTC TCCTTTTGAA AATGTAGAAT CTACAAACAA    2460

ACATCTAGAT GAAAATGAGA TTGTTGTTTC AATAAAGACA ACATCAAAAG ATAGGATGGG    2520

AAAAATGTTT ATAGATAAAA TTTTACTTGA AAGGTTTGTT AAACACCCTC AAAAAGTTAT    2580

AGGGATTTTC CTCAATGATG TACAAAGAAA AGAAGACAAC AATATCAGCT TTACACTTGT    2640

TTCAGGATTA TTTATGGTGT ATACTAAATT CTTAACTACT CTTGAAGGGA TCTATTATTT    2700

AGATCCACCA CCTAATGCAT TGAAACTACC ATATTCTAAT CATATGAAAA GATTTTCAGA    2760

TTTAATTACA GAAGACCTTG AAAAATTATT CTCCTCTTAA TTTTTTTCTT CTATTTTGAG    2820

AATTTGAATC ATCATCGCCT GAAAAAATCC TGTTGATCAT ACTTCGTCT TTCAATCGCT     2880

TTTGAATTTC CAGACAATAA TTTTTATCTA ACTCTGTTCC AATCCATTTT CTGCCAAAAG    2940

CTTCCGCAAC AGCATATGTA GTTCCCGAAC CACCAAATGG ATCAAGAATT AAACTATCTT    3000

TATTAGATGA AGATAAAATT ATTTTTTTAA GCAATTCTAC TGGCTTTTGC GTAGGGTGAA    3060

TACTTCTTTC CCAAGAACCA TTTGAAATTG TGGGAATCTC CAAAACATCT TTTGGCTTAG    3120
```

CTCCTAATGG GTTTGGCTCC CAAATATATT GCTTTTTTGA GTTCGAATAT TGCGAAGATT   3180

CGGCCTGGGT TCTTTGTGGA TATTTAACTG TATGCTCATT ATACGGGATT CGTGCCTCAT   3240

CAATATTAAA AATAAAATTT TTAGATTTTC TTAATAACAG TATACTTTCG TGTGAACGTC   3300

CCCAATCTTT ACCTAAATTT GCCTTGTTTC TATAGAACCA AATCAACCAT TTACAACTGT   3360

GAAAATATTT TGAAGTGATA AATTTTATGT CTGCCAGAAT TTCTGAAAAG CCACAAACAT   3420

ATAAACTGCC ATTATCTTTT AAAACTCTAT GAGCTTCCAT TATCCATTCT TTACTCCATT   3480

CGAGATATTC ATTTTGAGAA GAAAAAATAT CCCACTCTGC CTTTTTAATG TTATAAGGTG   3540

GATCTGCAAA AATTAAATCC ACAGAATTAC TCTCTAATTT TCTTAAAAAA TCAATAGCAT   3600

CAACATTGAA AAGTGAACCG TTTTTTGAAA TAAAAAATGG TTTAAGCTCG ATTTCCGAAT   3660

AAAGTTTTAA TTTTGGTATA TAACTGTAAT TATCAAACGC TATTTCATTC ACAAATGAAT   3720

CAATCTGCTG TTGTGTATAA ACCCTGTAAT TATTAATAGG ATGTCTTAAA CTTTTGAATT   3780

TTCCAGAATT ATCCCATCTT CCTTAATGTC TCAGAGTTAA CATCTAATAA TTTCGCCGCT   3840

TCTTTTATTG ATAAATAATC ATCCATATCT TACACAACAT TACACAAGTT TATACAGCAA   3900

ATATAAATAT TTTTTATACA TTGTAAAAAT TTTATTTACT TTTATTTTGT TCAATTGTCT   3960

CAATAAATAG TTAATCGAAA TACATTTTGA ATATGATAAA ATTGACTCCA ACAAATCTAA   4020

CACAATGACA TTAAAACCAA TAAAAACGGA AGAAGATTAC AATCAGGTTT TAGAAAGACT   4080

TTCACAAATT TTCGACGCTA AACCAAATAC CAAAGATGGA GATGAATTGG GAAATCTTGG   4140

GAATTC   4146

(2) INFORMATION FOR SEQ ID NO:30:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 23 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: DNA (genomic)

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATTTAGGTGA CACTATAGAA TAC   23

(2) INFORMATION FOR SEQ ID NO:31:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 49 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGCTCGAGT CTGATGACGA AGCTTGACTG ACTGAGATCA GCTTGCAAC   49

(2) INFORMATION FOR SEQ ID NO:32:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 35 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: double
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTCGAGATTT AGGTGACACT ATAGAATACA AGCTT   35

(2) INFORMATION FOR SEQ ID NO:33:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 60 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: double
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTCGAGATTT AGGTGACACT ATAGAATACG GAATTGTGAG CGGATAACAA TTCCAAGCTT   60